

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/SS0,671  
Source: pt 10  
Date Processed by STIC: 10/7/05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 10/07/2005

PATENT APPLICATION: US/10/550,671

TIME: 09:26:11

Input Set : A:\toyam114.007apc.txt

Output Set: N:\CRF4\10072005\J550671.raw

3 <110> APPLICANT: Yamaoka, Hideaki  
 4 Hoshijima, Mitsuhiro  
 5 Kawase, Shido  
 6 Kurosaka, Keisuke  
 8 <120> TITLE OF INVENTION: Method for producing glucose dehydrogenase  
 10 <130> FILE REFERENCE: TOYA114.007APC  
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/550,671  
 C--> 12 <141> CURRENT FILING DATE: 2005-09-26  
 12 <150> PRIOR APPLICATION NUMBER: PCT/JP2004/004074  
 13 <151> PRIOR FILING DATE: 2004-03-24  
 15 <150> PRIOR APPLICATION NUMBER: JP 2003-82739  
 16 <151> PRIOR FILING DATE: 2003-03-25  
 18 <160> NUMBER OF SEQ ID NOS: 15  
 20 <170> SOFTWARE: PatentIn Ver. 2.0  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 2467  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: Burkholderia cepacia  
 27 <220> FEATURE:  
 28 <221> NAME/KEY: CDS  
 29 <222> LOCATION: (258)..(761)  
 31 <220> FEATURE:  
 32 <221> NAME/KEY: CDS  
 33 <222> LOCATION: (764)..(2380)  
 35 <220> FEATURE:  
 36 <221> NAME/KEY: CDS  
 37 <222> LOCATION: (2386)..(2466)  
 39 <400> SEQUENCE: 1  
 40 aagcttttctg tttgattgca cgcgattcta accgagcgtc tgtgaggcgg aacgcgacat 60  
 41 gcttcgtgtc gcacacgtgt cgcgccgacg acacaaaaat gcagcgaaat ggctgatcgt 120  
 42 tacgaatggc tgacacattg aatggactat aaaaccattg tccgttccgg aatgtgcgcg 180  
 43 tacatttcag gtccgcgcgg atttttgaga aatatcaagc gtggttttcc cgaatccggt 240  
 44 gttcgagaga aggaaac atg cac aac gac aac act ccc cac tcg cgt cgc 290  
 45 Met His Asn Asp Asn Thr Pro His Ser Arg Arg  
 46 1 5 10  
 47 cac ggc gac gca gcc gca tca ggc atc acg cgg cgt caa tgg ttg caa 338  
 48 His Gly Asp Ala Ala Ser Gly Ile Thr Arg Arg Gln Trp Leu Gln  
 49 15 20 25  
 50 ggc gcg ctg gcg ctg acc gca gcg ggc ctc acg ggt tcg ctg aca ttg 386  
 51 Gly Ala Leu Ala Leu Thr Ala Ala Gly Leu Thr Gly Ser Leu Thr Leu  
 52 30 35 40  
 53 cgg gcg ctt gca gac aac ccc ggc act gcg ccg ctc gat acg ttc atg 434  
 54 Arg Ala Leu Ala Asp Asn Pro Gly Thr Ala Pro Leu Asp Thr Phe Met

P-6

## RAW SEQUENCE LISTING

DATE: 10/07/2005

PATENT APPLICATION: US/10/550,671

TIME: 09:26:11

Input Set : A:\toyam114.007apc.txt

Output Set: N:\CRF4\10072005\J550671.raw

55	45	50	55	
56	acg ctt tcc gaa tcg ctg acc ggc aag aaa ggg ctc agc cgc gtg atc	482		
57	Thr Leu Ser Glu Ser Leu Thr Gly Lys Lys Gly Leu Ser Arg Val Ile			
58	60	65	70	75
59	ggc gag cgc ctg ctg cag ggc ctg cag aag ggc tcg ttc aag acg gcc	530		
60	Gly Glu Arg Leu Leu Gln Ala Leu Gln Lys Gly Ser Phe Lys Thr Ala			
61		80	85	90
62	gac agc ctg ccg cag ctc gcc ggc gcg ctc gcg tcc ggt tcg ctg acg	578		
63	Asp Ser Leu Pro Gln Leu Ala Gly Ala Leu Ala Ser Gly Ser Leu Thr			
64		95	100	105
65	cct gaa cag gaa tcg ctc gca ctg acg atc ctc gag gcc tgg tat ctc	626		
66	Pro Glu Gln Glu Ser Leu Ala Leu Thr Ile Leu Glu Ala Trp Tyr Leu			
67		110	115	120
68	ggc atc gtc gac aac gtc gtg att acg tac gag gaa gca tta atg ttc	674		
69	Gly Ile Val Asp Asn Val Val Ile Thr Tyr Glu Glu Ala Leu Met Phe			
70		125	130	135
71	ggc gtc gtg tcc gat acg ctc gtg atc cgt tcg tat tgc ccc aac aaa	722		
72	Gly Val Val Ser Asp Thr Leu Val Ile Arg Ser Tyr Cys Pro Asn Lys			
73	140	145	150	155
74	ccc ggc ttc tgg gcc gac aaa ccg atc gag agg caa gcc tg atg gcc	769		
75	Pro Gly Phe Trp Ala Asp Lys Pro Ile Glu Arg Gln Ala Met Ala			
76		160	165	170
77	gat acc gat acg caa aag gcc gac gtc gtc gtc gtt gga tcg ggt gtc	817		
78	Asp Thr Asp Thr Gln Lys Ala Asp Val Val Val Val Gly Ser Gly Val			
79		175	180	185
80	gcg ggc gcg atc gtc gcg cat cag ctc gcg atg gcg ggc aag gcg gtg	865		
81	Ala Gly Ala Ile Val Ala His Gln Leu Ala Met Ala Gly Lys Ala Val			
82		190	195	200
83	atc ctg ctc gaa gcg ggc ccg cgc atg ccg cgc tgg gaa atc gtc gag	913		
84	Ile Leu Leu Glu Ala Gly Pro Arg Met Pro Arg Trp Glu Ile Val Glu			
85		205	210	215
86	cgc ttc cgc aat cag ccc gac aag atg gac ttc atg gcg ccg tac ccg	961		
87	Arg Phe Arg Asn Gln Pro Asp Lys Met Asp Phe Met Ala Pro Tyr Pro			
88		220	225	230
89	tcg agc ccc tgg gcg ccg cat ccc gag tac ggc ccg ccg aac gac tac	1009		
90	Ser Ser Pro Trp Ala Pro His Pro Glu Tyr Gly Pro Pro Asn Asp Tyr			
91	235	240	245	250
92	ctg atc ctg aag ggc gag cac aag ttc aac tcg cag tac atc cgc gcg	1057		
93	Leu Ile Leu Lys Gly Glu His Lys Phe Asn Ser Gln Tyr Ile Arg Ala			
94		255	260	265
95	gtg ggc ggc acg acg tgg cac tgg gcc gcg tcg gcg tgg cgc ttc att	1105		
96	Val Gly Gly Thr Thr Trp His Trp Ala Ala Ser Ala Trp Arg Phe Ile			
97		270	275	280
98	ccg aac gac ttc aag atg aag agc gtg tac ggc gtc ggc cgc gac tgg	1153		
99	Pro Asn Asp Phe Lys Met Lys Ser Val Tyr Gly Val Gly Arg Asp Trp			
100		285	290	295
101	ccg atc cag tac gac gat ctc gag ccg tac tat cag cgc gcg gag gaa	1201		
102	Pro Ile Gln Tyr Asp Asp Leu Glu Pro Tyr Tyr Gln Arg Ala Glu Glu			
103	300	305	310	

## RAW SEQUENCE LISTING

DATE: 10/07/2005

PATENT APPLICATION: US/10/550,671

TIME: 09:26:11

Input Set : A:\toyam114.007apc.txt

Output Set: N:\CRF4\10072005\J550671.raw

104	gag	ctc	ggc	gtg	tgg	ggc	ccg	ggc	ccc	gag	gaa	gat	ctg	tac	tcg	ccg	1249
105	Glu	Leu	Gly	Val	Trp	Gly	Pro	Gly	Pro	Glu	Glu	Asp	Leu	Tyr	Ser	Pro	
106	315					320				325						330	
107	cgc	aag	cag	ccg	tat	ccg	atg	ccg	ccg	ctg	ccg	ttg	tcg	ttc	aac	gag	1297
108	Arg	Lys	Gln	Pro	Tyr	Pro	Met	Pro	Pro	Leu	Pro	Leu	Ser	Phe	Asn	Glu	
109					335					340						345	
110	cag	acc	atc	aag	acg	gcg	ctg	aac	aac	tac	gat	ccg	aag	ttc	cat	gtc	1345
111	Gln	Thr	Ile	Lys	Thr	Ala	Leu	Asn	Asn	Tyr	Asp	Pro	Lys	Phe	His	Val	
112				350						355						360	
113	gtg	acc	gag	ccg	gtc	gcg	cgc	aac	agc	cgc	ccg	tac	gac	ggc	cgc	ccg	1393
114	Val	Thr	Glu	Pro	Val	Ala	Arg	Asn	Ser	Arg	Pro	Tyr	Asp	Gly	Arg	Pro	
115			365					370						375			
116	act	tgt	tgc	ggc	aac	aac	aac	tgc	atg	ccg	atc	tgc	ccg	atc	ggc	gcg	1441
117	Thr	Cys	Cys	Gly	Asn	Asn	Asn	Cys	Met	Pro	Ile	Cys	Pro	Ile	Gly	Ala	
118		380					385						390				
119	atg	tac	aac	ggc	atc	gtg	cac	gtc	gag	aag	gcc	gaa	cgc	gcc	ggc	gcg	1489
120	Met	Tyr	Asn	Gly	Ile	Val	His	Val	Glu	Lys	Ala	Glu	Arg	Ala	Gly	Ala	
121	395					400					405					410	
122	aag	ctg	atc	gag	aac	gcg	gtc	gtc	tac	aag	ctc	gag	acg	ggc	ccg	gac	1537
123	Lys	Leu	Ile	Glu	Asn	Ala	Val	Val	Tyr	Lys	Leu	Glu	Thr	Gly	Pro	Asp	
124					415					420						425	
125	aag	cgc	atc	gtc	gcg	gcg	ctc	tac	aag	gac	aag	acg	ggc	gcc	gag	cat	1585
126	Lys	Arg	Ile	Val	Ala	Ala	Leu	Tyr	Lys	Asp	Lys	Thr	Gly	Ala	Glu	His	
127			430						435						440		
128	cgc	gtc	gaa	ggc	aag	tat	ttc	gtg	ctc	gcc	gcg	aac	ggc	atc	gag	acg	1633
129	Arg	Val	Glu	Gly	Lys	Tyr	Phe	Val	Leu	Ala	Ala	Asn	Gly	Ile	Glu	Thr	
130			445					450						455			
131	ccg	aag	atc	ctg	ctg	atg	tcc	gcg	aac	cgc	gat	ttc	ccg	aac	ggt	gtc	1681
132	Pro	Lys	Ile	Leu	Leu	Met	Ser	Ala	Asn	Arg	Asp	Phe	Pro	Asn	Gly	Val	
133		460					465					470					
134	gcg	aac	agc	tcg	gac	atg	gtc	ggc	cgc	aac	ctg	atg	gac	cat	ccg	ggc	1729
135	Ala	Asn	Ser	Ser	Asp	Met	Val	Gly	Arg	Asn	Leu	Met	Asp	His	Pro	Gly	
136	475					480					485					490	
137	acc	ggc	gtg	tcg	ttc	tat	gcg	agc	gag	aag	ctg	tgg	ccg	ggc	cgc	ggc	1777
138	Thr	Gly	Val	Ser	Phe	Tyr	Ala	Ser	Glu	Lys	Leu	Trp	Pro	Gly	Arg	Gly	
139					495					500						505	
140	ccg	cag	gag	atg	acg	tcg	ctg	atc	ggt	ttc	cgc	gac	ggt	ccg	ttc	cgc	1825
141	Pro	Gln	Glu	Met	Thr	Ser	Leu	Ile	Gly	Phe	Arg	Asp	Gly	Pro	Phe	Arg	
142				510						515					520		
143	gcg	acc	gaa	gcg	gcg	aag	aag	atc	cac	ctg	tcg	aac	ctg	tcg	cgc	atc	1873
144	Ala	Thr	Glu	Ala	Ala	Lys	Lys	Ile	His	Leu	Ser	Asn	Leu	Ser	Arg	Ile	
145			525						530						535		
146	gac	cag	gag	acg	cag	aag	atc	ttc	aag	gcc	ggc	aag	ctg	atg	aag	ccc	1921
147	Asp	Gln	Glu	Thr	Gln	Lys	Ile	Phe	Lys	Ala	Gly	Lys	Leu	Met	Lys	Pro	
148		540						545							550		
149	gac	gag	ctc	gac	gcg	cag	atc	cgc	gac	cgt	tcc	gca	cgc	tac	gtg	cag	1969
150	Asp	Glu	Leu	Asp	Ala	Gln	Ile	Arg	Asp	Arg	Ser	Ala	Arg	Tyr	Val	Gln	
151	555					560					565					570	
152	ttc	gac	tgc	ttc	cac	gaa	atc	ctg	ccg	caa	ccc	gag	aac	cgc	atc	gtg	2017

## RAW SEQUENCE LISTING

DATE: 10/07/2005

PATENT APPLICATION: US/10/550,671

TIME: 09:26:11

Input Set : A:\toyam114.007apc.txt

Output Set: N:\CRF4\10072005\J550671.raw

```

153 Phe Asp Cys Phe His Glu Ile Leu Pro Gln Pro Glu Asn Arg Ile Val
154          575          580          585
155 ccg agc aag acg gcg acc gat gcg atc ggc att ccg cgc ccc gag atc 2065
156 Pro Ser Lys Thr Ala Thr Asp Ala Ile Gly Ile Pro Arg Pro Glu Ile
157          590          595          600
158 acg tat gcg atc gac gac tac gtg aag cgc ggc gcc gcg cat acg cgc 2113
159 Thr Tyr Ala Ile Asp Asp Tyr Val Lys Arg Gly Ala Ala His Thr Arg
160          605          610          615
161 gag gtc tac gcg acc gcc gcg aag gtg ctc ggc ggc acg gac gtc gtg 2161
162 Glu Val Tyr Ala Thr Ala Ala Lys Val Leu Gly Gly Thr Asp Val Val
163          620          625          630
164 ttc aac gac gaa ttc gcg ccg aac aat cac atc acg ggc tcg acg atc 2209
165 Phe Asn Asp Glu Phe Ala Pro Asn Asn His Ile Thr Gly Ser Thr Ile
166 635          640          645          650
167 atg ggc gcc gat gcg cgc gac tcc gtc gtc gac aag gac tgc cgc acg 2257
168 Met Gly Ala Asp Ala Arg Asp Ser Val Val Asp Lys Asp Cys Arg Thr
169          655          660          665
170 ttc gac cat ccg aac ctg ttc att tcg agc agc gcg acg atg ccg acc 2305
171 Phe Asp His Pro Asn Leu Phe Ile Ser Ser Ser Ala Thr Met Pro Thr
172          670          675          680
173 gtc ggt acc gta aac gtg acg ctg acg atc gcc gcg ctc gcg ctg cgg 2353
174 Val Gly Thr Val Asn Val Thr Leu Thr Ile Ala Ala Leu Ala Leu Arg
175          685          690          695
176 atg tcg gac acg ctg aag aag gaa gtc tgacc gtg cgg aaa tct act ctc 2403
177 Met Ser Asp Thr Leu Lys Lys Glu Val          Val Arg Lys Ser Thr Leu
178          700          705          710
179 act ttc ctc atc gcc ggc tgc ctc gcg ttg ccg ggc ttc gcg cgc gcg 2451
180 Thr Phe Leu Ile Ala Gly Cys Leu Ala Leu Pro Gly Phe Ala Arg Ala
181          715          720          725
182 gcc gat gcg gcc gat c 2467
183 Ala Asp Ala Ala Asp
184 730
186 <210> SEQ ID NO: 2
187 <211> LENGTH: 168
188 <212> TYPE: PRT
189 <213> ORGANISM: Burkholderia cepacia
191 <400> SEQUENCE: 2
192 Met His Asn Asp Asn Thr Pro His Ser Arg Arg His Gly Asp Ala Ala
193 1 5 10 15
194 Ala Ser Gly Ile Thr Arg Arg Gln Trp Leu Gln Gly Ala Leu Ala Leu
195 20 25 30
196 Thr Ala Ala Gly Leu Thr Gly Ser Leu Thr Leu Arg Ala Leu Ala Asp
197 35 40 45
198 Asn Pro Gly Thr Ala Pro Leu Asp Thr Phe Met Thr Leu Ser Glu Ser
199 50 55 60
200 Leu Thr Gly Lys Lys Gly Leu Ser Arg Val Ile Gly Glu Arg Leu Leu
201 65 70 75 80
202 Gln Ala Leu Gln Lys Gly Ser Phe Lys Thr Ala Asp Ser Leu Pro Gln
203 85 90 95

```

## RAW SEQUENCE LISTING

DATE: 10/07/2005

PATENT APPLICATION: US/10/550,671

TIME: 09:26:11

Input Set : A:\toyam114.007apc.txt

Output Set: N:\CRF4\10072005\J550671.raw

```

204 Leu Ala Gly Ala Leu Ala Ser Gly Ser Leu Thr Pro Glu Gln Glu Ser
205          100          105          110
206 Leu Ala Leu Thr Ile Leu Glu Ala Trp Tyr Leu Gly Ile Val Asp Asn
207          115          120          125
208 Val Val Ile Thr Tyr Glu Glu Ala Leu Met Phe Gly Val Val Ser Asp
209          130          135          140
210 Thr Leu Val Ile Arg Ser Tyr Cys Pro Asn Lys Pro Gly Phe Trp Ala
211 145          150          155          160
212 Asp Lys Pro Ile Glu Arg Gln Ala
213          165
214 <210> SEQ ID NO: 3
215 <211> LENGTH: 539
216 <212> TYPE: PRT
217 <213> ORGANISM: Burkholderia cepacia
219 <400> SEQUENCE: 3
220 Met Ala Asp Thr Asp Thr Gln Lys Ala Asp Val Val Val Val Gly Ser
221 1          5          10          15
222 Gly Val Ala Gly Ala Ile Val Ala His Gln Leu Ala Met Ala Gly Lys
223          20          25          30
224 Ala Val Ile Leu Leu Glu Ala Gly Pro Arg Met Pro Arg Trp Glu Ile
225          35          40          45
226 Val Glu Arg Phe Arg Asn Gln Pro Asp Lys Met Asp Phe Met Ala Pro
227          50          55          60
228 Tyr Pro Ser Ser Pro Trp Ala Pro His Pro Glu Tyr Gly Pro Pro Asn
229 65          70          75          80
230 Asp Tyr Leu Ile Leu Lys Gly Glu His Lys Phe Asn Ser Gln Tyr Ile
231          85          90          95
232 Arg Ala Val Gly Thr Thr Trp His Trp Ala Ala Ser Ala Trp Arg
233          100          105          110
234 Phe Ile Pro Asn Asp Phe Lys Met Lys Ser Val Tyr Gly Val Gly Arg
235          115          120          125
236 Asp Trp Pro Ile Gln Tyr Asp Asp Leu Glu Pro Tyr Tyr Gln Arg Ala
237          130          135          140
238 Glu Glu Glu Leu Gly Val Trp Gly Pro Gly Pro Glu Glu Asp Leu Tyr
239 145          150          155          160
240 Ser Pro Arg Lys Gln Pro Tyr Pro Met Pro Pro Leu Pro Leu Ser Phe
241          165          170          175
242 Asn Glu Gln Thr Ile Lys Thr Ala Leu Asn Asn Tyr Asp Pro Lys Phe
243          180          185          190
244 His Val Val Thr Glu Pro Val Ala Arg Asn Ser Arg Pro Tyr Asp Gly
245          195          200          205
246 Arg Pro Thr Cys Cys Gly Asn Asn Asn Cys Met Pro Ile Cys Pro Ile
247          210          215          220
248 Gly Ala Met Tyr Asn Gly Ile Val His Val Glu Lys Ala Glu Arg Ala
249 225          230          235          240
250 Gly Ala Lys Leu Ile Glu Asn Ala Val Val Tyr Lys Leu Glu Thr Gly
251          245          250          255
252 Pro Asp Lys Arg Ile Val Ala Ala Leu Tyr Lys Asp Lys Thr Gly Ala
253          260          265          270

```

## RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 10/07/2005

PATENT APPLICATION: US/10/550,671

TIME: 09:26:12

Input Set : A:\toyam114.007apc.txt

Output Set: N:\CRF4\10072005\J550671.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:6; Xaa Pos. 6,17,18,19,22

Seq#:11; Xaa Pos. 2,3

## VERIFICATION SUMMARY

DATE: 10/07/2005

PATENT APPLICATION: US/10/550,671

TIME: 09:26:12

Input Set : A:\toyam114.007apc.txt

Output Set: N:\CRF4\10072005\J550671.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:323 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0

M:341 Repeated in SeqNo=6

L:520 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0